

NOT  
RECEIVED  
21/12/1999

## AMENDED CLAIMS

[received by the International Bureau on 26 July 1999 (26.07.99);  
original claim 1 amended; original claims 2- 25 cancelled;  
new claims 2-88 added; (19 pages)]

4 What is claimed:

5 1. An invention substantially as described in the description.

6 2. An invention substantially as described and illustrated in the description.

7 3. A process for identifying one or more bi-allelic markers linked to a bi-allelic genetic characteristic  
8 gene in a species of creatures, comprising the steps of :  
9

10 a)choosing two or more bi-allelic covering markers so that a CL-F region is systematically covered by  
11 the two or more covering markers, the CL-F region being a collection of points on a two-dimensional  
12 plane, the two-dimensional plane having the two orthogonal dimensions of chromosomal location and  
13 least common allele frequency;

14  
15 b)choosing a statistical linkage test based on allelic association for each covering marker;

16  
17 c)choosing a sample of individuals for each covering marker ;  
18

19 d)obtaining genotype data/sample allele frequency data for each covering marker and the sample  
20 chosen for each covering marker, and obtaining phenotype status data for the genetic characteristic for  
21 each individual in the sample chosen for each covering marker;

22  
23 e)calculating evidence for linkage between each covering marker and the gene using the statistical  
24 linkage test based on allelic association chosen for each covering marker and the genotype  
25 data/sample allele frequency data for each covering marker and using the phenotype status data for the  
26 genetic characteristic for each individual in the sample chosen for each covering marker obtained in d);  
27 and  
28

29 f)identifying those covering markers as linked to the genetic characteristic gene which show evidence  
30 for linkage based on the calculations of step e.

31 4. A process as in claim 3, wherein the CL-F region is N covered to within a CL-F distance  $\delta$  by the two  
32 or more bi-allelic covering markers, so that each point in the region is within the CL-F distance  $\delta$  of N or  
33 more of the covering markers, wherein  $\delta$  is equal to about  $[\delta_{CL}, 0.25]$  or the equivalent thereof,  $\delta_{CL}$  is  
34 equal to the largest chromosomal length, computed by any method, for which linkage disequilibrium has  
35 been observed between any polymorphisms in any population of the species, N is an integer greater  
36 than or equal to 1.

37 5. A process as in claim 4, wherein the CL-F region comprises a CL-F matrix, the sum of the number of  
38 columns and rows in the matrix being greater than or equal to three, each cell of the matrix being of  
39 length  $L_{MC}$  and width  $W_{MC}$ , and  $L_{MC}$  being less than or equal to about  $\delta_{CL}$ , and  $W_{MC}$  being less than or  
40 equal to about 0.25,  $\delta_{CL}$  is equal to the largest chromosomal length, computed by any method, for which

1 linkage disequilibrium has been observed between any polymorphisms in any population of the species,  
2 there being N or more covering markers in each cell of the matrix and N is an integer greater than or  
3 equal to 1.

4 6. A process as in claim 5, wherein the covering markers are substantially nonevenly distributed across  
5 a chromosome or a chromosomal segment.

6 7. A process as in claim 5, wherein the covering markers are substantially evenly distributed across a  
7 chromosome or a chromosomal segment, and wherein the least common allele frequency of one or  
8 more markers is less than 0.4.

9 8. A process as in claim 5, wherein the covering markers are substantially evenly distributed across a  
10 chromosome or a chromosomal segment; and wherein there is a subgroup of one or more of the  
11 covering markers, and each of the markers in the subgroup is chosen without substantial preference for  
12 the least common allele frequency of each of the markers in the subgroup being close to 0.5.

13 9. A process as in claim 5, wherein the covering markers are substantially evenly distributed across a  
14 chromosome or a chromosomal segment, wherein (1) the average chromosomal intermarker distance  
15 of the covering markers is greater than 2 cM or the equivalent thereof and the least common allele  
16 frequency of one or more of the covering markers is less than 0.3 or wherein (2) the least common  
17 allele frequency of one or more of the covering markers is less than 0.2.

18 10. A process as in claim 5, wherein the covering markers are substantially evenly distributed across a  
19 chromosome or a chromosomal segment, wherein the average chromosomal intermarker distance of  
20 the covering markers is less than or equal to 2 cM or the equivalent thereof, and wherein the conditional  
21 probability the covering markers were chosen essentially randomly from substantially the known set of  
22 bi-allelic markers with least common allele frequencies between 0.2 inclusive and 0.5 inclusive is less  
23 than about 10 percent; wherein the conditional probability is substantially conditional on (1) the  
24 approximate chromosomal distribution of the covering markers, (2) the marker type of each covering  
25 marker and (3) there being N or more covering markers in each cell of the matrix.

26 11. A process as in claim 5, wherein the covering markers are substantially evenly distributed across a  
27 chromosome or a chromosomal segment, wherein the average chromosomal intermarker distance of  
28 the covering markers is greater than 2 cM or the equivalent thereof; and wherein the conditional  
29 probability the covering markers were chosen essentially randomly from substantially the known set of  
30 bi-allelic markers with least common allele frequencies between 0.3 inclusive and 0.5 inclusive is less  
31 than about 10 percent; wherein the conditional probability is substantially conditional on (1) the  
32 approximate chromosomal distribution of the covering markers, (2) the marker type of each covering  
33 marker and (3) there being N or more covering markers in each cell of the matrix.

34 12. A process as in claim 10, wherein the chromosome or the chromosomal segment consists  
35 essentially of a set of nonoverlapping chromosome segments of substantially equal length, and wherein  
36 one and only one covering marker is located on each of 80 percent or more of the chromosome  
37 segments of the set, and wherein zero or two and only two covering markers are located on each of 20  
38 percent or less of the chromosome segments of the set, and wherein each chromosome segment with  
39 zero covering markers located thereon is bordered only by chromosome segments with one or two  
40 covering markers located thereon, and wherein each chromosome segment with two covering markers

located thereon is bordered only by chromosome segments with one or zero covering markers located thereon; and wherein the conditional probability the covering markers were chosen essentially randomly from substantially the known set of bi-allelic markers with least common allele frequencies between 0.2 inclusive and 0.5 inclusive is less than about 10 percent; wherein the conditional probability is substantially conditional on (1) the chromosomal distribution of the covering markers on the chromosome segments of the set, (2) the marker type of each covering marker and (3) there being N or more covering markers in each cell of the matrix.

13. A process as in claim 11, wherein the chromosome or the chromosomal segment consists essentially of a set of nonoverlapping chromosome segments of substantially equal length, and wherein one and only one covering marker is located on each of 80 percent or more of the chromosome segments of the set, and wherein zero or two and only two covering markers are located on each of 20 percent or less of the chromosome segments of the set, and wherein each chromosome segment with zero covering markers located thereon is bordered only by chromosome segments with one or two covering markers located thereon, and wherein each chromosome segment with two covering markers located thereon is bordered only by chromosome segments with one or zero covering markers located thereon; and wherein the conditional probability the covering markers were chosen essentially randomly from substantially the known set of bi-allelic markers with least common allele frequencies between 0.3 inclusive and 0.5 inclusive is less than about 10 percent; wherein the conditional probability is substantially conditional on (1) the chromosomal distribution of the covering markers on the chromosome segments of the set, (2) the marker type of each covering marker and (3) there being N or more covering markers in each cell of the matrix.

14. A process as in claim 5, wherein the covering markers are substantially evenly distributed across a chromosome or a chromosomal segment, wherein the average chromosomal intermarker distance of the covering markers is less than or equal to 2 cM or the equivalent thereof, and wherein collection C is essentially the collection of known groups of bi-allelic markers with least common allele frequencies between 0.2 inclusive and 0.5 inclusive that are substantially similar to the covering markers as a group; wherein a group of bi-allelic markers is a member of collection C if and only if the group substantially meets criteria (1), (2), (3) and (4): (1) each marker in the group is chosen from substantially the known set of bi-allelic markers with least common allele frequencies between 0.2 inclusive and 0.5 inclusive, (2) the number of markers in the group is the same as the number of covering markers, (3) the chromosomal distribution of the group of markers and the covering markers is substantially similar, and (4) the marker type of each group marker and the covering marker with substantially the same chromosomal location is the same; wherein a group that is a member of collection C substantially meets criterion (5) if and only if (5) there are N or more of the group markers in each cell of the matrix; wherein P is essentially the proportion of groups in collection C that meet criterion (5); wherein P is less than about 90 percent.

15. A process as in claim 5, wherein the covering markers are substantially evenly distributed across a chromosome or a chromosomal segment, wherein the average chromosomal intermarker distance of the covering markers is greater than 2 cM or the equivalent thereof, and wherein collection C is essentially the collection of known groups of bi-allelic markers with least common allele frequencies

1 between 0.3 inclusive and 0.5 inclusive that are substantially similar to the covering markers as a  
2 group; wherein a group of bi-allelic markers is a member of collection C if and only if the group  
3 substantially meets criteria (1), (2), (3) and (4): (1) each marker in the group is chosen from  
4 substantially the known set of bi-allelic markers with least common allele frequencies between 0.3  
5 inclusive and 0.5 inclusive, (2) the number of markers in the group is the same as the number of  
6 covering markers, (3) the chromosomal distribution of the group of markers and the covering markers is  
7 substantially similar, and (4) the marker type of each group marker and the covering marker with  
8 substantially the same chromosomal location is the same; wherein a group that is a member of  
9 collection C substantially meets criterion (5) if and only if (5) there are N or more of the group markers  
10 in each cell of the matrix; wherein P is essentially the proportion of groups in collection C that meet  
11 criterion (5); wherein P is less than about 90 percent.

12 16. A process as in claim 5, wherein the covering markers are substantially evenly distributed across a  
13 chromosome or a chromosomal segment, wherein the average chromosomal intermarker distance of  
14 the covering markers is less than or equal to 2 cM or the equivalent thereof; wherein the chromosome  
15 or the chromosomal segment consists essentially of a set of nonoverlapping chromosome segments of  
16 substantially equal length, and wherein one and only one covering marker is located on each of 80  
17 percent or more of the chromosome segments of the set, and wherein zero or two and only two  
18 covering markers are located on each of 20 percent or less of the chromosome segments of the set,  
19 and wherein each chromosome segment with zero covering markers located thereon is bordered only  
20 by chromosome segments with one or two covering markers located thereon, and wherein each  
21 chromosome segment with two covering markers located thereon is bordered only by chromosome  
22 segments with one or zero covering markers located thereon; wherein collection D is essentially the  
23 collection of known groups of bi-allelic markers with least common allele frequencies between 0.2  
24 inclusive and 0.5 inclusive that are substantially similar to the covering markers as a group; wherein a  
25 group of bi-allelic markers is a member of collection D if and only if the group substantially meets  
26 criteria (1), (2), and (3): (1) each marker in the group is chosen from substantially the known set of bi-  
27 allelic markers with least common allele frequencies between 0.2 inclusive and 0.5 inclusive, (2) the  
28 number of covering markers and the number of group markers located on each chromosome segment  
29 of the set is the same, and (3) there is a group marker of the same type as each covering marker  
30 located on the same chromosome segment of the set as each covering marker; wherein a group that is  
31 a member of collection D substantially meets criterion (5) if and only if (5) there are N or more of the  
32 group markers in each cell of the matrix; wherein P is essentially the proportion of groups in collection D  
33 that meet criterion (5); wherein P is less than about 90 percent.

34 17. A process as in claim 5, wherein the covering markers are substantially evenly distributed across a  
35 chromosome or a chromosomal segment, wherein the average chromosomal intermarker distance of  
36 the covering markers is greater than 2 cM or the equivalent thereof; wherein the chromosome or the  
37 chromosomal segment consists essentially of a set of nonoverlapping chromosome segments of  
38 substantially equal length, and wherein one and only one covering marker is located on each of 80  
39 percent or more of the chromosome segments of the set, and wherein zero or two and only two  
40 covering markers are located on each of 20 percent or less of the chromosome segments of the set,

1 and wherein each chromosome segment with zero covering markers located thereon is bordered only  
2 by chromosome segments with one or two covering markers located thereon, and wherein each  
3 chromosome segment with two covering markers located thereon is bordered only by chromosome  
4 segments with one or zero covering markers located thereon; wherein collection D is essentially the  
5 collection of known groups of bi-allelic markers with least common allele frequencies between 0.3  
6 inclusive and 0.5 inclusive that are substantially similar to the covering markers as a group; wherein a  
7 group of bi-allelic markers is a member of collection D if and only if the group substantially meets  
8 criteria (1), (2), and (3): (1) each marker in the group is chosen from substantially the known set of bi-  
9 allelic markers with least common allele frequencies between 0.3 inclusive and 0.5 inclusive, (2) the  
10 number of covering markers and the number of group markers located on each chromosome segment  
11 of the set is the same, and (3) there is a group marker of the same type as each covering marker  
12 located on the same chromosome segment of the set as each covering marker; wherein a group that is  
13 a member of collection D substantially meets criterion (5) if and only if (5) there are N or more of the  
14 group markers in each cell of the matrix; wherein P is essentially the proportion of groups in collection D  
15 that meet criterion (5); wherein P is less than about 90 percent.

16 18. A process as in claim 4, wherein  $\delta$  is less than or equal to about [1 cM, 0.15] or the equivalent  
17 thereof.

18 19. A process as in claim 4, wherein (1) the covering markers are substantially nonevenly distributed  
19 across a chromosome or a chromosomal segment or (2) wherein the covering markers are substantially  
20 evenly distributed across a chromosome or a chromosomal segment, and wherein the least common  
21 allele frequency of one or more markers is less than 0.4 or (3) wherein the covering markers are  
22 substantially evenly distributed across a chromosome or a chromosomal segment; and wherein there is  
23 a subgroup of one or more of the covering markers, and each of the markers in the subgroup is chosen  
24 without substantial preference for the least common allele frequency of each of the markers in the  
25 subgroup being close to 0.5.

26 20. A process as in any one of claims 3-19, wherein there is a group of covering markers, and the  
27 markers in the group are a majority of the covering markers, and each marker in the group is an SNP,  
28 or a bi-allelic marker equivalent formed only from one or more SNPs.

29 21. A process as in claim 5 wherein  $L_{MC}$  is less than or equal to about 250,000 bp or the equivalent  
30 thereof,  $W_{MC}$  is less than or equal to about 0.15, wherein the species is human being, wherein the same  
31 statistical linkage test based on allelic association is chosen for each covering marker in step b) and  
32 wherein there is a group of covering markers, and the markers in the group are a majority of the  
33 covering markers, and each marker in the group is an SNP, or a bi-allelic marker equivalent formed  
34 only from one or more SNPs.

35 22. An apparatus for identifying bi-allelic markers linked to a bi-allelic genetic characteristic gene in a  
36 species of creatures, comprising: means to practice each of the steps of a process as in any one of the  
37 claims 3-21.

38 23. A process as in any one of claims 3-21, wherein the process comprises a computer program.

39 24. An apparatus as in claim 22, wherein the apparatus comprises a computer, the computer being  
40 supplied with proper data and instructions.

25. A process for localizing a bi-allelic genetic characteristic gene in a species of creatures to a CL-F location, comprising the steps of: any one of the processes in claims 3-21; further comprising: the step f) localizing the gene to the CL-F location of one or more markers that show evidence for linkage based on the calculations of step e).

26. A process as in claim 25, wherein the process comprises a computer program.

27. An apparatus for localizing a bi-allelic genetic characteristic gene in a species of creatures to a CL-F location, comprising: means to practice each of the steps of a process as in claim 25.

28. An apparatus for localizing a bi-allelic genetic characteristic gene in a species of creatures to a CL-F location as in claim 27, wherein the apparatus comprises a computer, the computer being supplied with proper data and instructions.

29. A process for obtaining genotype data/sample allele frequency data for each bi-allelic marker of a group of two or more bi-allelic covering markers in the chromosomal DNA of one or more individuals of a sample, each individual in the sample being a member of the same species, comprising:

a) determining information on the presence or absence of each allele of each bi-allelic marker of a group of two or more bi-allelic covering markers in the chromosomal DNA of one or more individuals of a sample, a CL-F region being systematically covered by the two or more bi-allelic covering markers, the CL-F region being a collection of points on a two-dimensional plane, the two-dimensional plane having the two orthogonal dimensions of chromosomal location and least common allele frequency; and

b) transforming the information of step a) into genotype data/sample allele frequency data for each marker of the group.

30. A process for obtaining genotype data/sample allele frequency data as in claim 29, wherein the CL-F region is N covered to within a CL-F distance  $\delta$  by the two or more bi-allelic covering markers, so that each point in the region is within the CL-F distance  $\delta$  of N or more of the covering markers, wherein  $\delta$  is equal to about  $[\delta_{CL}, 0.25]$  or the equivalent thereof,  $\delta_{CL}$  is equal to the largest chromosomal length, computed by any method, for which linkage disequilibrium has been observed between any polymorphisms in any population of the species, N is an integer greater than or equal to 1.

31. A process for obtaining genotype data/sample allele frequency data as in claim 30, wherein the CL-F region comprises a CL-F matrix, the sum of the number of columns and rows in the matrix being greater than or equal to three, each cell of the matrix being of length  $L_{MC}$  and width  $W_{MC}$ , and  $L_{MC}$  being less than or equal to about  $\delta_{CL}$ , and  $W_{MC}$  being less than or equal to about 0.25,  $\delta_{CL}$  is equal to the largest chromosomal length, computed by any method, for which linkage disequilibrium has been observed between any polymorphisms in any population of the species, there being N or more covering markers in each cell of the matrix and N is an integer greater than or equal to 1.

32. A process for obtaining genotype data/sample allele frequency data as in claim 31, wherein the covering markers are substantially nonevenly distributed across a chromosome or a chromosomal segment.

33. A process for obtaining genotype data/sample allele frequency data as in claim 31, wherein the covering markers are substantially evenly distributed across a chromosome or a chromosomal segment, and wherein the least common allele frequency of one or more markers is less than 0.4.

34. A process for obtaining genotype data/sample allele frequency data as in claim 31, wherein the covering markers are substantially evenly distributed across a chromosome or a chromosomal segment; and wherein there is a subgroup of one or more of the covering markers, and each of the markers in the subgroup is chosen without substantial preference for the least common allele frequency of each of the markers in the subgroup being close to 0.5.

35. A process for obtaining genotype data/sample allele frequency data as in claim 31, wherein the covering markers are substantially evenly distributed across a chromosome or a chromosomal segment, wherein (1) the average chromosomal intermarker distance of the covering markers is greater than 2 cM or the equivalent thereof and the least common allele frequency of one or more of the covering markers is less than 0.3 or wherein (2) the least common allele frequency of one or more of the covering markers is less than 0.2.

36. A process for obtaining genotype data/sample allele frequency data as in claim 31, wherein the covering markers are substantially evenly distributed across a chromosome or a chromosomal segment, wherein the average chromosomal intermarker distance of the covering markers is less than or equal to 2 cM or the equivalent thereof, and wherein the conditional probability the covering markers were chosen essentially randomly from substantially the known set of bi-allelic markers with least common allele frequencies between 0.2 inclusive and 0.5 inclusive is less than about 10 percent; wherein the conditional probability is substantially conditional on (1) the approximate chromosomal distribution of the covering markers, (2) the marker type of each covering marker and (3) there being N or more covering markers in each cell of the matrix.

37. A process for obtaining genotype data/sample allele frequency data as in claim 31, wherein the covering markers are substantially evenly distributed across a chromosome or a chromosomal segment, wherein the average chromosomal intermarker distance of the covering markers is greater than 2 cM or the equivalent thereof; and wherein the conditional probability the covering markers were chosen essentially randomly from substantially the known set of bi-allelic markers with least common allele frequencies between 0.3 inclusive and 0.5 inclusive is less than about 10 percent; wherein the conditional probability is substantially conditional on (1) the approximate chromosomal distribution of the covering markers, (2) the marker type of each covering marker and (3) there being N or more covering markers in each cell of the matrix.

38. A process for obtaining genotype data/sample allele frequency data as in claim 36, wherein the chromosome or the chromosomal segment consists essentially of a set of nonoverlapping chromosome segments of substantially equal length, and wherein one and only one covering marker is located on each of 80 percent or more of the chromosome segments of the set, and wherein zero or two and only two covering markers are located on each of 20 percent or less of the chromosome segments of the

1 set, and wherein each chromosome segment with zero covering markers located thereon is bordered  
2 only by chromosome segments with one or two covering markers located thereon, and wherein each  
3 chromosome segment with two covering markers located thereon is bordered only by chromosome  
4 segments with one or zero covering markers located thereon; and wherein the conditional probability  
5 the covering markers were chosen essentially randomly from substantially the known set of bi-allelic  
6 markers with least common allele frequencies between 0.2 inclusive and 0.5 inclusive is less than about  
7 10 percent; wherein the conditional probability is substantially conditional on (1) the chromosomal  
8 distribution of the covering markers on the chromosome segments of the set, (2) the marker type of  
9 each covering marker and (3) there being N or more covering markers in each cell of the matrix.

10 39. A process for obtaining genotype data/sample allele frequency data as in claim 37, wherein the  
11 chromosome or the chromosomal segment consists essentially of a set of nonoverlapping chromosome  
12 segments of substantially equal length, and wherein one and only one covering marker is located on  
13 each of 80 percent or more of the chromosome segments of the set, and wherein zero or two and only  
14 two covering markers are located on each of 20 percent or less of the chromosome segments of the  
15 set, and wherein each chromosome segment with zero covering markers located thereon is bordered  
16 only by chromosome segments with one or two covering markers located thereon, and wherein each  
17 chromosome segment with two covering markers located thereon is bordered only by chromosome  
18 segments with one or zero covering markers located thereon; and wherein the conditional probability  
19 the covering markers were chosen essentially randomly from substantially the known set of bi-allelic  
20 markers with least common allele frequencies between 0.3 inclusive and 0.5 inclusive is less than about  
21 10 percent; wherein the conditional probability is substantially conditional on (1) the chromosomal  
22 distribution of the covering markers on the chromosome segments of the set, (2) the marker type of  
23 each covering marker and (3) there being N or more covering markers in each cell of the matrix.

24 40. A process for obtaining genotype data/sample allele frequency data as in claim 31, wherein the  
25 covering markers are substantially evenly distributed across a chromosome or a chromosomal  
26 segment, wherein the average chromosomal intermarker distance of the covering markers is less than  
27 or equal to 2 cM or the equivalent thereof, and wherein collection C is essentially the collection of  
28 known groups of bi-allelic markers with least common allele frequencies between 0.2 inclusive and 0.5  
29 inclusive that are substantially similar to the covering markers as a group; wherein a group of bi-allelic  
30 markers is a member of collection C if and only if the group substantially meets criteria (1), (2), (3) and  
31 (4): (1) each marker in the group is chosen from substantially the known set of bi-allelic markers with  
32 least common allele frequencies between 0.2 inclusive and 0.5 inclusive, (2) the number of markers in  
33 the group is the same as the number of covering markers, (3) the chromosomal distribution of the group  
34 of markers and the covering markers is substantially similar, and (4) the marker type of each group  
35 marker and the covering marker with substantially the same chromosomal location is the same; wherein  
36 a group that is a member of collection C substantially meets criterion (5) if and only if (5) there are N or  
37 more of the group markers in each cell of the matrix; wherein P is essentially the proportion of groups in  
38 collection C that meet criterion (5); wherein P is less than about 90 percent.

39 41. A process for obtaining genotype data/sample allele frequency data as in claim 31, wherein the  
40 covering markers are substantially evenly distributed across a chromosome or a chromosomal



segment, wherein the average chromosomal intermarker distance of the covering markers is greater than 2 cM or the equivalent thereof, and wherein collection C is essentially the collection of known groups of bi-allelic markers with least common allele frequencies between 0.3 inclusive and 0.5 inclusive that are substantially similar to the covering markers as a group; wherein a group of bi-allelic markers is a member of collection C if and only if the group substantially meets criteria (1), (2), (3) and (4): (1) each marker in the group is chosen from substantially the known set of bi-allelic markers with least common allele frequencies between 0.3 inclusive and 0.5 inclusive, (2) the number of markers in the group is the same as the number of covering markers, (3) the chromosomal distribution of the group of markers and the covering markers is substantially similar, and (4) the marker type of each group marker and the covering marker with substantially the same chromosomal location is the same; wherein a group that is a member of collection C substantially meets criterion (5) if and only if (5) there are N or more of the group markers in each cell of the matrix; wherein P is essentially the proportion of groups in collection C that meet criterion (5); wherein P is less than about 90 percent.

42. A process for obtaining genotype data/sample allele frequency data as in claim 31, wherein the covering markers are substantially evenly distributed across a chromosome or a chromosomal segment, wherein the average chromosomal intermarker distance of the covering markers is less than or equal to 2 cM or the equivalent thereof; wherein the chromosome or the chromosomal segment consists essentially of a set of nonoverlapping chromosome segments of substantially equal length, and wherein one and only one covering marker is located on each of 80 percent or more of the chromosome segments of the set, and wherein zero or two and only two covering markers are located on each of 20 percent or less of the chromosome segments of the set, and wherein each chromosome segment with zero covering markers located thereon is bordered only by chromosome segments with one or two covering markers located thereon, and wherein each chromosome segment with two covering markers located thereon is bordered only by chromosome segments with one or zero covering markers located thereon; wherein collection D is essentially the collection of known groups of bi-allelic markers with least common allele frequencies between 0.2 inclusive and 0.5 inclusive that are substantially similar to the covering markers as a group; wherein a group of bi-allelic markers is a member of collection D if and only if the group substantially meets criteria (1), (2), and (3): (1) each marker in the group is chosen from substantially the known set of bi-allelic markers with least common allele frequencies between 0.2 inclusive and 0.5 inclusive, (2) the number of covering markers and the number of group markers located on each chromosome segment of the set is the same, and (3) there is a group marker of the same type as each covering marker located on the same chromosome segment of the set as each covering marker; wherein a group that is a member of collection D substantially meets criterion (5) if and only if (5) there are N or more of the group markers in each cell of the matrix; wherein P is essentially the proportion of groups in collection D that meet criterion (5); wherein P is less than about 90 percent.

43. A process for obtaining genotype data/sample allele frequency data as in claim 31, wherein the covering markers are substantially evenly distributed across a chromosome or a chromosomal segment, wherein the average chromosomal intermarker distance of the covering markers is greater than 2 cM or the equivalent thereof; wherein the chromosome or the chromosomal segment consists

essentially of a set of nonoverlapping chromosome segments of substantially equal length, and wherein one and only one covering marker is located on each of 80 percent or more of the chromosome segments of the set, and wherein zero or two and only two covering markers are located on each of 20 percent or less of the chromosome segments of the set, and wherein each chromosome segment with zero covering markers located thereon is bordered only by chromosome segments with one or two covering markers located thereon, and wherein each chromosome segment with two covering markers located thereon is bordered only by chromosome segments with one or zero covering markers located thereon; wherein collection D is essentially the collection of known groups of bi-allelic markers with least common allele frequencies between 0.3 inclusive and 0.5 inclusive that are substantially similar to the covering markers as a group; wherein a group of bi-allelic markers is a member of collection D if and only if the group substantially meets criteria (1), (2), and (3): (1) each marker in the group is chosen from substantially the known set of bi-allelic markers with least common allele frequencies between 0.3 inclusive and 0.5 inclusive, (2) the number of covering markers and the number of group markers located on each chromosome segment of the set is the same, and (3) there is a group marker of the same type as each covering marker located on the same chromosome segment of the set as each covering marker; wherein a group that is a member of collection D substantially meets criterion (5) if and only if (5) there are N or more of the group markers in each cell of the matrix; wherein P is essentially the proportion of groups in collection D that meet criterion (5); wherein P is less than about 90 percent.

44. A process for obtaining genotype data/sample allele frequency data as in claim 30, wherein  $\delta$  is less than or equal to about [1 cM, 0.15] or the equivalent thereof.

45. A process for obtaining genotype data/sample allele frequency data as in claim 30, wherein (1) the covering markers are substantially nonevenly distributed across a chromosome or a chromosomal segment or (2) wherein the covering markers are substantially evenly distributed across a chromosome or a chromosomal segment, and wherein the least common allele frequency of one or more markers is less than 0.4 or (3) wherein the covering markers are substantially evenly distributed across a chromosome or a chromosomal segment; and wherein there is a subgroup of one or more of the covering markers, and each of the markers in the subgroup is chosen without substantial preference for the least common allele frequency of each of the markers in the subgroup being close to 0.5.

46. A process for obtaining genotype data/sample allele frequency data as in any one of claims 29-45, wherein there is a group of covering markers, and the markers in the group are a majority of the covering markers, and each marker in the group is an SNP, or a bi-allelic marker equivalent formed only from one or more SNPs.

47. A process for obtaining genotype data/sample allele frequency data as in claim 31 wherein  $L_{MC}$  is less than or equal to about 250,000 bp or the equivalent thereof,  $W_{MC}$  is less than or equal to about 0.15, wherein the species is human being, wherein the same statistical linkage test based on allelic association is chosen for each covering marker in step b) and wherein there is a group of covering markers, and the markers in the group are a majority of the covering markers, and each marker in the group is an SNP, or a bi-allelic marker equivalent formed only from one or more SNPs.

1 48. An apparatus for obtaining genotype data/sample allele frequency data for each bi-allelic marker of  
2 a group of two or more bi-allelic covering markers in the chromosomal DNA of one or more individuals  
3 of a sample, each individual in the sample being a member of the same species, comprising: means to  
4 practice each of the steps of a process as in any one of the claims 29-47.

5 49. A process for obtaining genotype data/sample allele frequency data as in any one of claims 29-47,  
6 wherein the process comprises a computer program.

7 50. An apparatus as in claim 48, wherein the apparatus comprises a computer, the computer being  
8 supplied with proper data and instructions.

✓  
10 51. The use of one or more copies of a set of oligonucleotides to determine genotype data/sample  
11 allele frequency data for each bi-allelic marker of a group of two or more bi-allelic covering markers for  
12 one or more individuals, each individual being a member of the same species, wherein the group of  
13 covering markers systematically cover a CL-F region, the CL-F region being a collection of points on a  
14 two-dimensional plane, the two-dimensional plane having the two orthogonal dimensions of  
15 chromosomal location and least common allele frequency.

16 52. The use as in claim 51, wherein the CL-F region is N covered to within a CL-F distance  $\delta$  by the two  
17 or more bi-allelic covering markers, so that each point in the region is within the CL-F distance  $\delta$  of N or  
18 more of the covering markers, wherein  $\delta$  is equal to about  $[\delta_{CL}, 0.25]$  or the equivalent thereof,  $\delta_{CL}$  is  
19 equal to the largest chromosomal length, computed by any method, for which linkage disequilibrium has  
20 been observed between any polymorphisms in any population of the species, N is an integer greater  
21 than or equal to 1.

22 53. The use as in claim 51, wherein the CL-F region comprises a CL-F matrix, the sum of the number of  
23 columns and rows in the matrix being greater than or equal to three, each cell of the matrix being of  
24 length  $L_{MC}$  and width  $W_{MC}$ , and  $L_{MC}$  being less than or equal to about  $\delta_{CL}$ , and  $W_{MC}$  being less than or  
25 equal to about 0.25,  $\delta_{CL}$  is equal to the largest chromosomal length, computed by any method, for which  
26 linkage disequilibrium has been observed between any polymorphisms in any population of the species,  
27 there being N or more covering markers in each cell of the matrix and N is an integer greater than or  
28 equal to 1.

29 54. The use as in claim 53, wherein the covering markers are substantially nonevenly distributed across  
30 a chromosome or a chromosomal segment.

31 55. The use as in claim 53, wherein the covering markers are substantially evenly distributed across a  
32 chromosome or a chromosomal segment, and wherein the least common allele frequency of one or  
33 more markers is less than 0.4.

34 56. The use as in claim 53, wherein the covering markers are substantially evenly distributed across a  
35 chromosome or a chromosomal segment; and wherein there is a subgroup of one or more of the  
36 covering markers, and each of the markers in the subgroup is chosen without substantial preference for  
37 the least common allele frequency of each of the markers in the subgroup being close to 0.5.

38 57. The use as in claim 53, wherein the covering markers are substantially evenly distributed across a  
39 chromosome or a chromosomal segment, wherein (1) the average chromosomal intermarker distance  
40 of the covering markers is greater than 2 cM or the equivalent thereof and the least common allele

1 frequency of one or more of the covering markers is less than 0.3 or wherein (2) the least common  
2 allele frequency of one or more of the covering markers is less than 0.2.

3 58. The use as in claim 53, wherein the covering markers are substantially evenly distributed across a  
4 chromosome or a chromosomal segment, wherein the average chromosomal intermarker distance of  
5 the covering markers is less than or equal to 2 cM or the equivalent thereof, and wherein the conditional  
6 probability the covering markers were chosen essentially randomly from substantially the known set of  
7 bi-allelic markers with least common allele frequencies between 0.2 inclusive and 0.5 inclusive is less  
8 than about 10 percent; wherein the conditional probability is substantially conditional on (1) the  
9 approximate chromosomal distribution of the covering markers, (2) the marker type of each covering  
10 marker and (3) there being N or more covering markers in each cell of the matrix.

11 59. The use as in claim 53, wherein the covering markers are substantially evenly distributed across a  
12 chromosome or a chromosomal segment, wherein the average chromosomal intermarker distance of  
13 the covering markers is greater than 2 cM or the equivalent thereof; and wherein the conditional  
14 probability the covering markers were chosen essentially randomly from substantially the known set of  
15 bi-allelic markers with least common allele frequencies between 0.3 inclusive and 0.5 inclusive is less  
16 than about 10 percent; wherein the conditional probability is substantially conditional on (1) the  
17 approximate chromosomal distribution of the covering markers, (2) the marker type of each covering  
18 marker and (3) there being N or more covering markers in each cell of the matrix.

19 60. The use as in claim 58, wherein the chromosome or the chromosomal segment consists essentially  
20 of a set of nonoverlapping chromosome segments of substantially equal length, and wherein one and  
21 only one covering marker is located on each of 80 percent or more of the chromosome segments of the  
22 set, and wherein zero or two and only two covering markers are located on each of 20 percent or less  
23 of the chromosome segments of the set, and wherein each chromosome segment with zero covering  
24 markers located thereon is bordered only by chromosome segments with one or two covering markers  
25 located thereon, and wherein each chromosome segment with two covering markers located thereon is  
26 bordered only by chromosome segments with one or zero covering markers located thereon; and  
27 wherein the conditional probability the covering markers were chosen essentially randomly from  
28 substantially the known set of bi-allelic markers with least common allele frequencies between 0.2  
29 inclusive and 0.5 inclusive is less than about 10 percent; wherein the conditional probability is  
30 substantially conditional on (1) the chromosomal distribution of the covering markers on the  
31 chromosome segments of the set, (2) the marker type of each covering marker and (3) there being N or  
32 more covering markers in each cell of the matrix.

33 61. The use as in claim 59, wherein the chromosome or the chromosomal segment consists essentially  
34 of a set of nonoverlapping chromosome segments of substantially equal length, and wherein one and  
35 only one covering marker is located on each of 80 percent or more of the chromosome segments of the  
36 set, and wherein zero or two and only two covering markers are located on each of 20 percent or less  
37 of the chromosome segments of the set, and wherein each chromosome segment with zero covering  
38 markers located thereon is bordered only by chromosome segments with one or two covering markers  
39 located thereon, and wherein each chromosome segment with two covering markers located thereon is  
40 bordered only by chromosome segments with one or zero covering markers located thereon; and

wherein the conditional probability the covering markers were chosen essentially randomly from substantially the known set of bi-allelic markers with least common allele frequencies between 0.3 inclusive and 0.5 inclusive is less than about 10 percent; wherein the conditional probability is substantially conditional on (1) the chromosomal distribution of the covering markers on the chromosome segments of the set, (2) the marker type of each covering marker and (3) there being N or more covering markers in each cell of the matrix.

62. The use as in claim 53, wherein the covering markers are substantially evenly distributed across a chromosome or a chromosomal segment, wherein the average chromosomal intermarker distance of the covering markers is less than or equal to 2 cM or the equivalent thereof, and wherein collection C is essentially the collection of known groups of bi-allelic markers with least common allele frequencies between 0.2 inclusive and 0.5 inclusive that are substantially similar to the covering markers as a group; wherein a group of bi-allelic markers is a member of collection C if and only if the group substantially meets criteria (1), (2), (3) and (4): (1) each marker in the group is chosen from substantially the known set of bi-allelic markers with least common allele frequencies between 0.2 inclusive and 0.5 inclusive, (2) the number of markers in the group is the same as the number of covering markers, (3) the chromosomal distribution of the group of markers and the covering markers is substantially similar, and (4) the marker type of each group marker and the covering marker with substantially the same chromosomal location is the same; wherein a group that is a member of collection C substantially meets criterion (5) if and only if (5) there are N or more of the group markers in each cell of the matrix; wherein P is essentially the proportion of groups in collection C that meet criterion (5); wherein P is less than about 90 percent.

63. The use as in claim 53, wherein the covering markers are substantially evenly distributed across a chromosome or a chromosomal segment, wherein the average chromosomal intermarker distance of the covering markers is greater than 2 cM or the equivalent thereof, and wherein collection C is essentially the collection of known groups of bi-allelic markers with least common allele frequencies between 0.3 inclusive and 0.5 inclusive that are substantially similar to the covering markers as a group; wherein a group of bi-allelic markers is a member of collection C if and only if the group substantially meets criteria (1), (2), (3) and (4): (1) each marker in the group is chosen from substantially the known set of bi-allelic markers with least common allele frequencies between 0.3 inclusive and 0.5 inclusive, (2) the number of markers in the group is the same as the number of covering markers, (3) the chromosomal distribution of the group of markers and the covering markers is substantially similar, and (4) the marker type of each group marker and the covering marker with substantially the same chromosomal location is the same; wherein a group that is a member of collection C substantially meets criterion (5) if and only if (5) there are N or more of the group markers in each cell of the matrix; wherein P is essentially the proportion of groups in collection C that meet criterion (5); wherein P is less than about 90 percent.

64. The use as in claim 53, wherein the covering markers are substantially evenly distributed across a chromosome or a chromosomal segment, wherein the average chromosomal intermarker distance of the covering markers is less than or equal to 2 cM or the equivalent thereof; wherein the chromosome or the chromosomal segment consists essentially of a set of nonoverlapping chromosome segments of

1 substantially equal length, and wherein one and only one covering marker is located on each of 80  
2 percent or more of the chromosome segments of the set, and wherein zero or two and only two  
3 covering markers are located on each of 20 percent or less of the chromosome segments of the set,  
4 and wherein each chromosome segment with zero covering markers located thereon is bordered only  
5 by chromosome segments with one or two covering markers located thereon, and wherein each  
6 chromosome segment with two covering markers located thereon is bordered only by chromosome  
7 segments with one or zero covering markers located thereon; wherein collection D is essentially the  
8 collection of known groups of bi-allelic markers with least common allele frequencies between 0.2  
9 inclusive and 0.5 inclusive that are substantially similar to the covering markers as a group; wherein a  
10 group of bi-allelic markers is a member of collection D if and only if the group substantially meets  
11 criteria (1), (2), and (3): (1) each marker in the group is chosen from substantially the known set of bi-  
12 allelic markers with least common allele frequencies between 0.2 inclusive and 0.5 inclusive, (2) the  
13 number of covering markers and the number of group markers located on each chromosome segment  
14 of the set is the same, and (3) there is a group marker of the same type as each covering marker  
15 located on the same chromosome segment of the set as each covering marker; wherein a group that is  
16 a member of collection D substantially meets criterion (5) if and only if (5) there are N or more of the  
17 group markers in each cell of the matrix; wherein P is essentially the proportion of groups in collection D  
18 that meet criterion (5); wherein P is less than about 90 percent.

19 65. The use as in claim 53, wherein the covering markers are substantially evenly distributed across a  
20 chromosome or a chromosomal segment, wherein the average chromosomal intermarker distance of  
21 the covering markers is greater than 2 cM or the equivalent thereof; wherein the chromosome or the  
22 chromosomal segment consists essentially of a set of nonoverlapping chromosome segments of  
23 substantially equal length, and wherein one and only one covering marker is located on each of 80  
24 percent or more of the chromosome segments of the set, and wherein zero or two and only two  
25 covering markers are located on each of 20 percent or less of the chromosome segments of the set,  
26 and wherein each chromosome segment with zero covering markers located thereon is bordered only  
27 by chromosome segments with one or two covering markers located thereon, and wherein each  
28 chromosome segment with two covering markers located thereon is bordered only by chromosome  
29 segments with one or zero covering markers located thereon; wherein collection D is essentially the  
30 collection of known groups of bi-allelic markers with least common allele frequencies between 0.3  
31 inclusive and 0.5 inclusive that are substantially similar to the covering markers as a group; wherein a  
32 group of bi-allelic markers is a member of collection D if and only if the group substantially meets  
33 criteria (1), (2), and (3): (1) each marker in the group is chosen from substantially the known set of bi-  
34 allelic markers with least common allele frequencies between 0.3 inclusive and 0.5 inclusive, (2) the  
35 number of covering markers and the number of group markers located on each chromosome segment  
36 of the set is the same, and (3) there is a group marker of the same type as each covering marker  
37 located on the same chromosome segment of the set as each covering marker; wherein a group that is  
38 a member of collection D substantially meets criterion (5) if and only if (5) there are N or more of the  
39 group markers in each cell of the matrix; wherein P is essentially the proportion of groups in collection D  
40 that meet criterion (5); wherein P is less than about 90 percent.

1 66. The use as in claim 52, wherein  $\delta$  is less than or equal to about [1 cM, 0.15] or the equivalent  
2 thereof.

3 67. The use as in claim 52, wherein (1) the covering markers are substantially nonevenly distributed  
4 across a chromosome or a chromosomal segment or (2) wherein the covering markers are substantially  
5 evenly distributed across a chromosome or a chromosomal segment, and wherein the least common  
6 allele frequency of one or more markers is less than 0.4 or (3) wherein the covering markers are  
7 substantially evenly distributed across a chromosome or a chromosomal segment; and wherein there is  
8 a subgroup of one or more of the covering markers, and each of the markers in the subgroup is chosen  
9 without substantial preference for the least common allele frequency of each of the markers in the  
10 subgroup being close to 0.5.

11 68. The use as in any one of claims 51-67, wherein there is a group of covering markers, and the  
12 markers in the group are a majority of the covering markers, and each marker in the group is an SNP,  
13 or a bi-allelic marker equivalent formed only from one or more SNPs.

14 69. The use as in claim 53 wherein  $L_{MC}$  is less than or equal to about 250,000 bp or the equivalent  
15 thereof,  $W_{MC}$  is less than or equal to about 0.15, wherein the species is human being, wherein the same  
16 statistical linkage test based on allelic association is chosen for each covering marker in step b) and  
17 wherein there is a group of covering markers, and the markers in the group are a majority of the  
18 covering markers, and each marker in the group is an SNP, or a bi-allelic marker equivalent formed  
19 only from one or more SNPs.

20  
21 70. One or more copies of a set of oligonucleotides, the set of oligonucleotides being complementary to  
22 a group of two or more bi-allelic covering markers of the same species, wherein the group of covering  
23 markers systematically cover a CL-F region, the CL-F region being a collection of points on a two-  
24 dimensional plane, the two-dimensional plane having the two orthogonal dimensions of chromosomal  
25 location and least common allele frequency.

26 71. One or more copies of a set of oligonucleotides as in claim 70, wherein the CL-F region is N  
27 covered to within a CL-F distance  $\delta$  by the two or more bi-allelic covering markers, so that each point in  
28 the region is within the CL-F distance  $\delta$  of N or more of the covering markers, wherein  $\delta$  is equal to  
29 about [ $\delta_{CL}$ , 0.25] or the equivalent thereof,  $\delta_{CL}$  is equal to the largest chromosomal length, computed by  
30 any method, for which linkage disequilibrium has been observed between any polymorphisms in any  
31 population of the species, N is an integer greater than or equal to 1.

32 72. One or more copies of a set of oligonucleotides as in claim 70, wherein the CL-F region comprises  
33 a CL-F matrix, the sum of the number of columns and rows in the matrix being greater than or equal to  
34 three, each cell of the matrix being of length  $L_{MC}$  and width  $W_{MC}$ , and  $L_{MC}$  being less than or equal to  
35 about  $\delta_{CL}$ , and  $W_{MC}$  being less than or equal to about 0.25,  $\delta_{CL}$  is equal to the largest chromosomal  
36 length, computed by any method, for which linkage disequilibrium has been observed between any  
37 polymorphisms in any population of the species, there being N or more covering markers in each cell of  
38 the matrix and N is an integer greater than or equal to 1.

39 73. One or more copies of a set of oligonucleotides as in claim 72, wherein the covering markers are  
40 substantially nonevenly distributed across a chromosome or a chromosomal segment.

- 1 74. One or more copies of a set of oligonucleotides as in claim 72, wherein the covering markers are  
2 substantially evenly distributed across a chromosome or a chromosomal segment, and wherein the  
3 least common allele frequency of one or more markers is less than 0.4.
- 4 75. One or more copies of a set of oligonucleotides as in claim 72, wherein the covering markers are  
5 substantially evenly distributed across a chromosome or a chromosomal segment; and wherein there is  
6 a subgroup of one or more of the covering markers, and each of the markers in the subgroup is chosen  
7 without substantial preference for the least common allele frequency of each of the markers in the  
8 subgroup being close to 0.5.
- 9 76. One or more copies of a set of oligonucleotides as in claim 72, wherein the covering markers are  
10 substantially evenly distributed across a chromosome or a chromosomal segment, wherein (1) the  
11 average chromosomal intermarker distance of the covering markers is greater than 2 cM or the  
12 equivalent thereof and the least common allele frequency of one or more of the covering markers is  
13 less than 0.3 or wherein (2) the least common allele frequency of one or more of the covering markers  
14 is less than 0.2.
- 15 77. One or more copies of a set of oligonucleotides as in claim 72, wherein the covering markers are  
16 substantially evenly distributed across a chromosome or a chromosomal segment, wherein the average  
17 chromosomal intermarker distance of the covering markers is less than or equal to 2 cM or the  
18 equivalent thereof, and wherein the conditional probability the covering markers were chosen  
19 essentially randomly from substantially the known set of bi-allelic markers with least common allele  
20 frequencies between 0.2 inclusive and 0.5 inclusive is less than about 10 percent; wherein the  
21 conditional probability is substantially conditional on (1) the approximate chromosomal distribution of  
22 the covering markers, (2) the marker type of each covering marker and (3) there being N or more  
23 covering markers in each cell of the matrix.
- 24 78. One or more copies of a set of oligonucleotides as in claim 72, wherein the covering markers are  
25 substantially evenly distributed across a chromosome or a chromosomal segment, wherein the average  
26 chromosomal intermarker distance of the covering markers is greater than 2 cM or the equivalent  
27 thereof; and wherein the conditional probability the covering markers were chosen essentially randomly  
28 from substantially the known set of bi-allelic markers with least common allele frequencies between 0.3  
29 inclusive and 0.5 inclusive is less than about 10 percent; wherein the conditional probability is  
30 substantially conditional on (1) the approximate chromosomal distribution of the covering markers, (2)  
31 the marker type of each covering marker and (3) there being N or more covering markers in each cell of  
32 the matrix.
- 33 79. One or more copies of a set of oligonucleotides as in claim 77, wherein the chromosome or the  
34 chromosomal segment consists essentially of a set of nonoverlapping chromosome segments of  
35 substantially equal length, and wherein one and only one covering marker is located on each of 80  
36 percent or more of the chromosome segments of the set, and wherein zero or two and only two  
37 covering markers are located on each of 20 percent or less of the chromosome segments of the set,  
38 and wherein each chromosome segment with zero covering markers located thereon is bordered only  
39 by chromosome segments with one or two covering markers located thereon, and wherein each  
40 chromosome segment with two covering markers located thereon is bordered only by chromosome



segments with one or zero covering markers located thereon; and wherein the conditional probability the covering markers were chosen essentially randomly from substantially the known set of bi-allelic markers with least common allele frequencies between 0.2 inclusive and 0.5 inclusive is less than about 10 percent; wherein the conditional probability is substantially conditional on (1) the chromosomal distribution of the covering markers on the chromosome segments of the set, (2) the marker type of each covering marker and (3) there being N or more covering markers in each cell of the matrix.

80. One or more copies of a set of oligonucleotides as in claim 78, wherein the chromosome or the chromosomal segment consists essentially of a set of nonoverlapping chromosome segments of substantially equal length, and wherein one and only one covering marker is located on each of 80 percent or more of the chromosome segments of the set, and wherein zero or two and only two covering markers are located on each of 20 percent or less of the chromosome segments of the set, and wherein each chromosome segment with zero covering markers located thereon is bordered only by chromosome segments with one or two covering markers located thereon, and wherein each chromosome segment with two covering markers located thereon is bordered only by chromosome segments with one or zero covering markers located thereon; and wherein the conditional probability the covering markers were chosen essentially randomly from substantially the known set of bi-allelic markers with least common allele frequencies between 0.3 inclusive and 0.5 inclusive is less than about 10 percent; wherein the conditional probability is substantially conditional on (1) the chromosomal distribution of the covering markers on the chromosome segments of the set, (2) the marker type of each covering marker and (3) there being N or more covering markers in each cell of the matrix.

81. One or more copies of a set of oligonucleotides as in claim 72, wherein the covering markers are substantially evenly distributed across a chromosome or a chromosomal segment, wherein the average chromosomal intermarker distance of the covering markers is less than or equal to 2 cM or the equivalent thereof, and wherein collection C is essentially the collection of known groups of bi-allelic markers with least common allele frequencies between 0.2 inclusive and 0.5 inclusive that are substantially similar to the covering markers as a group; wherein a group of bi-allelic markers is a member of collection C if and only if the group substantially meets criteria (1), (2), (3) and (4): (1) each marker in the group is chosen from substantially the known set of bi-allelic markers with least common allele frequencies between 0.2 inclusive and 0.5 inclusive, (2) the number of markers in the group is the same as the number of covering markers, (3) the chromosomal distribution of the group of markers and the covering markers is substantially similar, and (4) the marker type of each group marker and the covering marker with substantially the same chromosomal location is the same; wherein a group that is a member of collection C substantially meets criterion (5) if and only if (5) there are N or more of the group markers in each cell of the matrix; wherein P is essentially the proportion of groups in collection C that meet criterion (5); wherein P is less than about 90 percent.

82. One or more copies of a set of oligonucleotides as in claim 72, wherein the covering markers are substantially evenly distributed across a chromosome or a chromosomal segment, wherein the average chromosomal intermarker distance of the covering markers is greater than 2 cM or the equivalent thereof, and wherein collection C is essentially the collection of known groups of bi-allelic markers with least common allele frequencies between 0.3 inclusive and 0.5 inclusive that are substantially similar to

the covering markers as a group; wherein a group of bi-allelic markers is a member of collection C if and only if the group substantially meets criteria (1), (2), (3) and (4): (1) each marker in the group is chosen from substantially the known set of bi-allelic markers with least common allele frequencies between 0.3 inclusive and 0.5 inclusive, (2) the number of markers in the group is the same as the number of covering markers, (3) the chromosomal distribution of the group of markers and the covering markers is substantially similar, and (4) the marker type of each group marker and the covering marker with substantially the same chromosomal location is the same; wherein a group that is a member of collection C substantially meets criterion (5) if and only if (5) there are N or more of the group markers in each cell of the matrix; wherein P is essentially the proportion of groups in collection C that meet criterion (5); wherein P is less than about 90 percent.

83. One or more copies of a set of oligonucleotides as in claim 72, wherein the covering markers are substantially evenly distributed across a chromosome or a chromosomal segment, wherein the average chromosomal intermarker distance of the covering markers is less than or equal to 2 cM or the equivalent thereof; wherein the chromosome or the chromosomal segment consists essentially of a set of nonoverlapping chromosome segments of substantially equal length, and wherein one and only one covering marker is located on each of 80 percent or more of the chromosome segments of the set, and wherein zero or two and only two covering markers are located on each of 20 percent or less of the chromosome segments of the set, and wherein each chromosome segment with zero covering markers located thereon is bordered only by chromosome segments with one or two covering markers located thereon, and wherein each chromosome segment with two covering markers located thereon is bordered only by chromosome segments with one or zero covering markers located thereon; wherein collection D is essentially the collection of known groups of bi-allelic markers with least common allele frequencies between 0.2 inclusive and 0.5 inclusive that are substantially similar to the covering markers as a group; wherein a group of bi-allelic markers is a member of collection D if and only if the group substantially meets criteria (1), (2), and (3): (1) each marker in the group is chosen from substantially the known set of bi-allelic markers with least common allele frequencies between 0.2 inclusive and 0.5 inclusive, (2) the number of covering markers and the number of group markers located on each chromosome segment of the set is the same, and (3) there is a group marker of the same type as each covering marker located on the same chromosome segment of the set as each covering marker; wherein a group that is a member of collection D substantially meets criterion (5) if and only if (5) there are N or more of the group markers in each cell of the matrix; wherein P is essentially the proportion of groups in collection D that meet criterion (5); wherein P is less than about 90 percent.

84. One or more copies of a set of oligonucleotides as in claim 72, wherein the covering markers are substantially evenly distributed across a chromosome or a chromosomal segment, wherein the average chromosomal intermarker distance of the covering markers is greater than 2 cM or the equivalent thereof; wherein the chromosome or the chromosomal segment consists essentially of a set of nonoverlapping chromosome segments of substantially equal length, and wherein one and only one covering marker is located on each of 80 percent or more of the chromosome segments of the set, and wherein zero or two and only two covering markers are located on each of 20 percent or less of the

1 chromosome segments of the set, and wherein each chromosome segment with zero covering markers  
2 located thereon is bordered only by chromosome segments with one or two covering markers located  
3 thereon, and wherein each chromosome segment with two covering markers located thereon is  
4 bordered only by chromosome segments with one or zero covering markers located thereon; wherein  
5 collection D is essentially the collection of known groups of bi-allelic markers with least common allele  
6 frequencies between 0.3 inclusive and 0.5 inclusive that are substantially similar to the covering  
7 markers as a group; wherein a group of bi-allelic markers is a member of collection D if and only if the  
8 group substantially meets criteria (1), (2), and (3): (1) each marker in the group is chosen from  
9 substantially the known set of bi-allelic markers with least common allele frequencies between 0.3  
10 inclusive and 0.5 inclusive, (2) the number of covering markers and the number of group markers  
11 located on each chromosome segment of the set is the same, and (3) there is a group marker of the  
12 same type as each covering marker located on the same chromosome segment of the set as each  
13 covering marker; wherein a group that is a member of collection D substantially meets criterion (5) if  
14 and only if (5) there are N or more of the group markers in each cell of the matrix; wherein P is  
15 essentially the proportion of groups in collection D that meet criterion (5); wherein P is less than about  
16 90 percent.

17 85. One or more copies of a set of oligonucleotides as in claim 71, wherein  $\delta$  is less than or equal to  
18 about [1 cM, 0.15] or the equivalent thereof.

19 86. One or more copies of a set of oligonucleotides as in claim 71, wherein (1) the covering markers are  
20 substantially nonevenly distributed across a chromosome or a chromosomal segment or (2) wherein the  
21 covering markers are substantially evenly distributed across a chromosome or a chromosomal  
22 segment, and wherein the least common allele frequency of one or more markers is less than 0.4 or (3)  
23 wherein the covering markers are substantially evenly distributed across a chromosome or a  
24 chromosomal segment; and wherein there is a subgroup of one or more of the covering markers, and  
25 each of the markers in the subgroup is chosen without substantial preference for the least common  
26 allele frequency of each of the markers in the subgroup being close to 0.5.

27 87. One or more copies of a set of oligonucleotides as in any one of claims 70-86, wherein there is a  
28 group of covering markers, and the markers in the group are a majority of the covering markers, and  
29 each marker in the group is an SNP, or a bi-allelic marker equivalent formed only from one or more  
30 SNPs.

31 88. One or more copies of a set of oligonucleotides as in claim 72 wherein  $L_{MC}$  is less than or equal to  
32 about 250,000 bp or the equivalent thereof,  $W_{MC}$  is less than or equal to about 0.15, wherein the  
33 species is human being, wherein the same statistical linkage test based on allelic association is chosen  
34 for each covering marker in step b) and wherein there is a group of covering markers, and the markers  
35 in the group are a majority of the covering markers, and each marker in the group is an SNP, or a bi-  
36 allelic marker equivalent formed only from one or more SNPs.

## Statement under Article 19(1)

Some of the amended claims make use of the phrase "conditional probability", such as claim 11. Some of the amended claims make use of the phrase "proportion of groups", such as claim 14. There are various techniques to calculate or estimate such a probability or such a proportion. These techniques include, but are not necessarily limited to, direct calculation, statistical estimates, and Monte Carlo estimation techniques. Powerful software is available for calculation and statistical estimation for data in matrix format or two-dimensional format. Some such software is available from Cytel Software Corporation, Cambridge, Massachusetts ( example: Exact Logistic Regression: Theory and Examples, Mehta CR, Patel NR, Statistics in Medicine, vol 14, 2143-2160(1995). Another example is SAS (SAS Institute Inc., SAS Campus Drive, Cary, North Carolina 27513, USA.; A handbook of statistical analyses using SAS by Brian S. Everitt and Geoff Der, Boca Raton, Fla. : Chapman & Hall/CRC, 1998.). A further example is MATLAB (The MathWorks, Inc. 3 Apple Hill Drive, Natick, Mass. U.S.A. 01760-2098; MATLAB primer by Kermit Sigmon, 4th ed. Boca Raton : CRC Press, c1994.) Statistical techniques include techniques for hypothesis testing, goodness-of-fit and others.

The degree of skill in the art in probability and statistics is great. Indeed the inventor's important equation (Equation 2, page 38) is an equation for  $P_t$ , wherein  $P_t$  is a binomial probability for parental allele 'transmission' which determines the magnitude of the TDT chi-square statistic.  $P_s$  (pages 40-42) is also a binomial probability that determines the magnitude of the ASP test statistic. (see Abstract and Paper: Annals of Human Genetics (1998), 62, 159-179. The abstract is available on the World Wide Web and Internet, including at the journal's website.) Skill in the use of computers in the art is also great (page 25).

Some claims, such as claims 11, 12, 13, 14 and others make use of the phrase "substantially the known set of bi-allelic markers". As pointed out in the description (page 25) information on bi-allelic markers can be gained from sources such as the Whitehead Institute or Marshfield Foundation for Biomedical Research. Similar sources of information on Single Nucleotide Polymorphisms can be obtained from sources given in SNP attack on complex traits, Nature Genetics, volume 20 no. 3, Nov 1998, pp. 217-218.

Some claims, such as claims 11, 12, 13, 14 and others make use of the term "marker type" or similar terminology. As stated in the description, a bi-allelic marker may be an SNP, a microsatellite marker, a bi-allelic marker equivalent formed from one or more true bi-allelic markers. "Marker type" means type of true bi-allelic marker as for example an SNP or a microsatellite; or "marker type" means a bi-allelic marker equivalent of a certain type, such as a bi-allelic marker equivalent formed only from one or more SNPs or a bi-allelic marker equivalent formed only from one or more microsatellites.)